DOCUMENT RESUME

ED 437 332 SO 031 372

AUTHOR Grant, S. G.
TITLE Teachers and Tests: Elementary and Secondary School Teachers' Perceptions of Changes in the New York State Testing Program.
PUB DATE 1999-11-00
NOTE 38p.; Paper presented at the Annual Meeting of the National Council for the Social Studies (79th, Orlando, FL, November 19-21, 1999).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Elementary Secondary Education; *Faculty Development; *Focus Groups; *Standardized Tests; *State Programs; Student Evaluation; *Teacher Attitudes; *Testing Programs
IDENTIFIERS Interpretive Research; *New York; Relationship Change Research

ABSTRACT

This paper uses data collected through focus group interviews over two years to explore the relationship between teachers and tests. In the first year of data collection, two focus groups were organized; one group was composed of 7 elementary school teachers and counselors and the other group was composed of 12 high school teachers. The participants, with varied teaching experiences, represented a cross section of urban, suburban, and rural school districts in western New York State. Each of the 2-hour focus group interviews was tape recorded and transcribed. During the second year of data collection, separate elementary and secondary focus groups were organized. Post-interview evaluations were completed by participants. All data were analyzed from an interpretive stance emphasizing the importance of context and the multiple ways individuals construct meaning. The three patterns synthesized from the data relate to the nature and substance of the tests, the professional development opportunities available to teachers, and the rationales for and the consequences of the state examinations. Analysis suggests that teachers see the New York State tests as a mixed bag. Teachers across grade levels and subject areas express an uneasy combination of hope and fear, anticipation and dread. The focus group teachers reported that school-, grade-, and/or department-level professional development activities generally were more useful than state or district efforts. Findings suggest that teachers need to be much more involved in the process of changing state assessments and professional development needs to be more attuned to the different needs teachers have. (Contains 48 references; focus group protocol is appended.) (BT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 437 332

Teachers and Tests: Elementary and Secondary School Teachers' Perceptions of Changes in the New York State Testing Program

S. G. Grant
517 Baldy Hall
State University of New York at Buffalo
Buffalo, New York 14260
Telephone: 716-645-6493
Email: sggrant@acsu.buffalo.edu

SO 031 372

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

S.G. Grant

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paper presented at the annual conference of the College and University Faculty Association of the National Council for the Social Studies, Orlando, FL.
November, 1999

BEST COPY AVAILABLE

Teachers and Tests:
Elementary and Secondary School Teachers' Perceptions of Changes in the New York State Testing Program

Many current initiatives seek to raise educational standards and improve student academic performance. Yet, there is a curious gap in the recent talk about national and state reforms. While much attention focuses on defining higher expectations for what students will know and be able to do, little attention is given to how teachers should learn new pedagogical ideas and practices. Such policies as the federal *Goals 2000: Educate America Act* and the New York *New Compact for Learning* focus on the resources, conditions, and practices necessary for all students to learn. None of these efforts, however, seriously addresses the education profession, that is, how experienced teachers will learn the intended innovations (National Center for Education Statistics, 1999).

How do teachers change their pedagogical practices? Some suggest change comes through new subject matter standards proposed by professional organizations (National Council for Social Studies, 1994), by national groups (National Center for History in the Schools, 1994), or by state education departments (New York State Education Department, 1996). Others believe teachers change their practices in response to organizational restructuring (e.g., smaller classes, block scheduling). Still others assert that real change in the classroom lives of teachers and students depends on changes in state-level assessments (Comfort, 1991; Smith & O'Day, 1991). The assumption here is that testing drives much of what teachers do, and so curricular and instructional change will occur if and when state tests change.

This last idea is intriguing for, if true, it suggests the potential for big pedagogical changes with a modicum of policy effort: Change the test and one changes teachers' practices. New York state policymakers seem taken with this approach, for although they have developed new curriculum standards, it is revision of the state testing program which gets most of the attention (Grant, 1997a). The scope of that revision is wide. One piece is the change from program evaluation tests at the elementary level to high-stakes individual student testing. A second piece is the phase-out of the less demanding high school Regents Competency Tests and the requirement

that all students pass the more demanding Regents tests. A third piece is a change in the content and format of all state tests presumably to reflect the higher expectations expressed in the state's new standards documents.

What sense do teachers make of these new state tests and how, if at all, do the tests influence their classroom practices? Strange as it seems, there is little empirical evidence to suggest how teachers, especially teachers at different grade levels, respond to changes in state tests. Assessment is a particularly hot topic in educational circles today, yet there is surprisingly little research which digs deeply into teachers' understandings of the import of standardized tests (Cohen & Barnes, 1993; Grant, 1998b). Corbett and Wilson's (1991) study of teachers' reactions to a new Maryland testing program is well-known, but it is one of few such studies in a field that is more prone to study students' responses than teachers'.

In this paper, I use the data collected through focus group interviews over two years to explore the relationships between teachers and tests. My findings suggest that teachers need to be much more involved in the process of changing state assessments, and that professional development needs to be more attuned to the different needs teachers have.

The Study

The Teacher Learning and Assessment (TLA) research project is designed to look generally at the intersection of teachers and assessments. More specifically, we are interested in exploring the relationship between teacher learning and state-level testing. Our study questions include: a) in what ways are tests and test results used in classrooms, schools, and the districts; b) what do the proposed changes in state-level tests mean for teachers and learners; c) how are teachers being prepared to respond to the new state assessments; and d) what challenges do teachers and administrators anticipate in moving toward new state assessments? In each case, we are interested in if and how these issues differ across school subject matters and grade levels.

Data Collection

In the first year of data collection, we organized two focus groups, one composed of 7 elementary school teachers and counselors and one composed of 12 high school teachers. The

participants represented a cross-section of urban, suburban, and rural school districts in western New York state, a breadth of teaching experience (2-25 years), and a range of school subjects (language arts, mathematics, science, and social studies). Each of the two-hour focus group interviews was tape-recorded and transcribed.

During the second year of data collection, we again organized separate elementary and secondary focus groups. As we did, we debated whether to a) reconstitute the original groups only; b) develop new groups of teachers separate from those involved in the first year's interviews; or c) call together groups that mixed teachers new to the project with those who had participated during the previous year. We rejected the first option, fearing that attrition might leave us with groups that were too small. We also rejected the second option, though largely because of timing: We did not think we could hold four focus groups near the end of the school year. In the end, we decided to constitute mixed groups for two reasons. One reason was that we wanted to expand the number of teachers we were talking with; the second reason is that we were interested in how the two groups might interact. The secondary focus group consisted of 8 teachers representing mathematics, science, English, and social studies; 5 of the 8 were in the original sample. The elementary focus group consisted of 5 teachers, 3 of whom were in the original sample.

The data consist of interview transcripts of the focus group sessions and post-interview evaluations completed by the participants. The focus group interviews follow a semi-structured interview protocol (see Appendix). Questions used during the first year asked participants to construct a metaphor to represent their sense of the changes in state-level testing, what the new tests mean for teaching and learning across school subjects, how teachers are being prepared for new standards and new assessments, and what challenges teachers believe they face. The post-interview questions asked the participants to reflect on the issues raised around the relationship between state-level assessment and classroom practice. The interview protocol was largely the same during year two. Changes consisted of replacing the metaphor task with a fill-in-the-blank exercise ("I used to think of the state assessment as _____, now I [still] think of it as

_____.") and the addition of probes that asked participants if they sensed a change from last year to the present. There were no changes to the post-interview evaluation.

Data Analysis

All data were analyzed inductively from an interpretivist stance (Bogdan & Biklen, 1982; LeCompte, Preissle, & Tesch, 1993). That stance emphasizes the importance of context, and the multiple ways individuals construct meaning. All data were also analyzed using a constant comparative method (Bogdan & Biklen, 1982; Glaser, 1978). That method assumes that data collection and analysis are recursive, one informing the other throughout the course of the study. After coding the data both within and across grade levels and subject matters, I began seeking patterns in the informants' responses. The themes which emerged reflect the full data set, but in each case I highlight the implications for social studies.

Interview and evaluation data were analyzed for patterns and themes related to the research questions. In the analysis of the focus group interviews, particular attention was given to issues of how teachers make sense of, and make different sense of, the reform documents they encounter; the kinds of learning opportunities they attend, and how, if at all, these reforms and opportunities influence what teachers think about and do in their classrooms. Looking across the interviews, I saw patterns which help explain the teachers' responses in a social context and the nature of their learning in an array of social settings. The three patterns I synthesized from the data and report on in this paper relate to the nature and substance of the tests, the professional development opportunities available to teachers, and the rationales for and the consequences of the state exams.

On Tests and Teaching

Standardized tests matter. The professional literature is replete with debates about tests as a means of accountability, as measures of performance, and as levers of change (Corbett & Wilson, 1991; Editors, 1994; Feltovich, Spiro, & Coulson, 1993; Finn, 1995; Fuhrman, Clune, & Elmore, 1988; Koretz, 1988; Ravitch, 1995; Resnick & Resnick, 1985). These concerns become elevated when situations like CTB/McGraw-Hill's mis-scoring of almost 9000 New York City students' tests occur. In all of the talk about tests, however, one area gets scant regard: What teachers learn

from tests, and if and how that knowledge affects their instructional practice. Common sense holds that tests drive classroom instruction. Evidence for that opinion is thin, however. Much research focuses on the relationship between students and tests (See, for example, Natriello & Pallas, 1998; Stiggins & Conklin, 1992; Wolf, 1998), but few empirical studies explore the relationship between teachers and the tests they administer (Cohen & Barnes, 1993; Corbett & Wilson, 1991; Firestone, Mayrowetz, & Fairman, 1998; Grant, 1998b). The research that is available presents a mixed picture at best.

Those advocates of tests as a vehicle for driving educational change tend to cite general positive effects rather than specifics. Some (Feltovich et al., 1993; Popham, 1998; Shanker, 1995) simply argue that good tests will inevitably drive good instruction. Lacking any more specificity, Popham, Cruse, Rankin, Sandifer, and Williams (1985) claim that tests measure important learning, and that good tests results equal good education. Systemic reformers (Fuhrman, 1993; Smith & O'Day, 1991) advocate for testing as part of an overall strategy aimed at fundamental school change. Others (English, 1980; Glatthorn, 1987; Heubert & Hauser, 1999) argue that because standardized tests are a reality in most school districts, they should be used as a fundamental part of curriculum planning.

Critics of standardized testing are more direct in their assessment of the impact of testing on teaching. Madaus (1988) claims, among other things, that teachers will teach to the test, that they will adjust their instruction to follow the form of the questions asked (e.g., multiple-choice, essay), and that tests transfer control over the curriculum to whomever controls the test.[1] Claims by LeMahieu (1984) and Koretz (1995) are more tentative, but they too conclude that teachers may tailor their curricula to the content covered on the test. Recent empirical work supports some of these claims. Firestone, Mayrowetz, and Fairman (1998) assert that testing programs in Maine and Maryland seem to influence teachers' content decisions, although they conclude that such influences are weaker than expected. Corbett and Wilson (1991) argue that testing, especially minimum-competency testing, has a pernicious effect on teachers in that it causes them to narrow their sense of educational purposes and to focus on activities designed to raise test scores whether

or not they think those activities are good for students. They conclude that squeezing teachers in this fashion encourages them to rebel against reform measures good and bad. "Statewide testing programs do control activity at the local level, but the subsequent activity is not reform" (p. 1).

Other researchers are less sure that a direct relationship exists between standardized testing and teachers' classroom practices. Freeman, Kuhs, Porter, Knappen, Floden, Schmidt, & Schwille (1980), Kelleghan, Madaus, and Airasian (1982), and Salmon-Cox (1981) found little direct impact of standardized testing on teachers' daily instruction. Firestone, Mayrowetz, and Fairman (1998) claim that, while tests may have influenced teachers' decisions about what to teach, there was virtually no influence on their decisions about how to teach. In a cross-case comparison of two high school teachers' civil rights units (Grant, 1998b), I found little direct influence of testing on either teacher's content or pedagogical decision-making.

This brief review suggests two points. First, we need to know more about the relationship between teachers and tests. While the impact of tests on students has been much explored, research that inquires into if, how, and in what ways teachers are influenced by standardized tests is lacking. Second, that research around teachers and tests fails to show a clear or consistent pattern of influence. Tests matter, but how and to what extent is unclear.

State-Level Curriculum and Assessment in New York State

State-level influence over curriculum and assessment is a well-established tradition in New York State. The Regents test has been administered continually for over 100 years. These tests are administered in all academic subjects and are tied to school courses. For example, in social studies, students take the Global Studies test at the end of a two-year Global Studies course sequence in ninth and tenth grades; eleventh graders take the U. S. History and Government test after completing a course of the same name. Elementary and middle school teachers also follow a state curriculum in all school subjects and students take state-developed tests.

Recent State-Level Curriculum Changes

As is the case in most states, educational reform has been steady work since the 1980s. Begun during the tenure of former Commissioner of Education, Thomas Sobol, state-level activity

around school curriculum hit full stride in the mid-1990s under current Commissioner Richard Mills.

Since 1994, working groups of state policymakers, teachers, and administrators have produced new curriculum and learning standards and scope and sequences for all school subjects. Social studies teachers, for example, may now consult the *Learning Standards for Social Studies* (New York State Education Department, 1996) and the *Resource Guide for Social Studies* (New York State Education Department, 1998). Compared with the previous round of curricular revisions in the mid-to-late 1980s, the changes represented in these documents vary from virtually no change in the K-5 grades curricula, which follows an expanding horizons model, in the seventh and eighth grade U.S. and New York State history, or in the twelfth grade Participation in Government and Economics courses. Modest changes are evident in other curricula, such as the emphasis on geography in the eleventh grade U.S. history and government course. Major changes seem localized at sixth grade, where the course of study expanded from Western and Eastern Europe and the Middle East to the entire Eastern hemisphere, and at ninth and tenth grades, where the emphasis has changed from a cultural approach as represented in Global Studies to a chronological study as expressed as Global History and Geography.

Recent State-Level Assessment Changes

The state-level testing program is also changing. Although the scope of the changes varies[2], the net effect appears to be a general ratcheting up of the stakes for both teacher and students.

State tests of language arts, mathematics, and science have undergone radical transformations which include reducing the number of multiple-choice items and increasing the number and range of performance tasks. For example, new science tests call for students to actually perform experiments. By contrast, the social studies assessments will apparently change little: Multiple-choice questions will still dominate the tests, accounting for 55% of a student's score.[3] The major change seems to be in the writing portion of the exam. Unlike many minimum competency tests, New York students have always had to answer essay questions on state exams. The new tests are different primarily in the fact that a) students will no longer have a range of essay prompts to

choose from, and b) a new kind of essay question, a document-based question (DBQ), is being introduced on each of the fifth, eighth, tenth, and eleventh grade tests. A DBQ asks students to write an essay synthesizing a number of primary source documents (e.g., short quotes from government documents and famous individuals, political cartoons, poems, charts and graphs).[4] Plans call for students to answer a main idea-type question about each of the documents before writing their essay. High school students will also write a second, "thematic" essay based on a single prompt.[5] The inclusion of the DBQ is the primary change in the structure of the social studies exams. One might argue that such a question represents a major shift away from traditional testing, but given the scope of the test (and the fact that students can easily pass the test without a single DBQ point), adding a DBQ could be read as a minor revision.

Three other changes seem more dramatic. One is that the new fifth and eighth grade tests will produce individual student scores. Tests at those levels, termed "Program Evaluation Tests," have aimed at helping teachers understand the effectiveness of their content and pedagogical decisions.[6] The shift of emphasis to individual students is apparently intended to raise the stakes of these tests and tie them more directly to the high school Regents exams. The function of the Regents test is also being fundamentally changed. In the past, passing Regents tests in all academic subjects meant that a student earned a Regents diploma. Students could opt to take the less rigorous Regents Competency Exam (RCT) and earn a local diploma. Ninth graders beginning in 2001 will no longer have these options. The RCT will no longer be administered, and all students will have to pass five Regents examinations (English, mathematics, global history, U.S. history, and science) in order to graduate.

Given these changes, state-level tests are no less high-stakes for teachers than they are for students. Since the mid-1990s, state policymakers have introduced a number of curriculum reforms, such as new state standards for social studies, yet it is concern about the state tests which surfaces most regularly in teachers' talk (Grant, 1997a). This makes sense for two reasons. First, the curriculum documents produced thus far offer teachers little assistance in making concrete instructional decisions (Grant, 1997b). Second, the messages teachers receive often promote the

view that tests are intended to drive change (Grant, 1996). For example, during sessions devoted to new state social studies standards, one representative from the New York State Education Department (NYSED) said that new tests will "help grow change in the system." During another session, a different SED representative said, "New assessments will represent a change in instruction....Kids won't perform well until (teachers') instruction reflects this." And at yet a third meeting, NYSED Commissioner Richard Mills added, "Instruction won't change until the tests change." The message that tests matter was echoed during local school and district meetings. A suburban district social studies supervisor, for example, told teachers that "change in content will come if we change the tests." An urban district supervisor observed, "If we change the assessments, we'll change instruction" (p. 271). One might question the focus of test influence--instruction, curriculum, or the "system" in general--but it is hard to miss the larger point: tests matter.

The Prospects and Problems of State-Level Testing In New York State

The tendency of advocates and critics to cast standardized testing in black and white images is not supported here. My analysis suggests that teachers see the new NYS tests as a mixed bag. The prospects of tests which more closely mirror and support thoughtful instruction and closer collaboration with colleagues are mitigated by the problems of, among other things, uncertainty about the rationale for and consequences of the new tests and the uneveness of the opportunities to learn about and respond to changes in the tests. In short, teachers across grade levels and subject matters express an uneasy combination of hope and fear, anticipation and dread. I explore those poles by looking at teachers' perceptions of the new tests in terms of their nature and substance, the professional development opportunities available, and the rationales and consequences.

The Nature and Substance of the New NYS Tests

The NYSED is phasing in the new state tests over a period of four years, beginning with the English language arts tests at grade 4 in January, 1999. Consequently most of the teachers interviewed have not seen final versions of the tests they will administer. All have, however, received preliminary materials from state, district, and professional organization sources and so

most assume that they have a fair sense of what the new exams will be like. Most believe the tests will be an improvement over past assessments, but questions about the nature and substance arise.

Both elementary and secondary teachers expressed at least modest support for the general direction taken in the new tests. A middle school science teacher suggested simply that the NYSED was "changing what assessment means." An elementary school teacher was more specific. "I think there was a lot of change going on and then they changed the assessment," she said, "I remember giving that CTBS (a basic skills test) and teaching a literature-based program, and we were all complaining that it wasn't reflective [of our teaching]." Another elementary school teacher was more specific: "The new assessments test the same way we teach reading, and where we want kids to be in math."

Social studies teachers approved of the move to include primary sources within the DBQ. A high school teacher cited the real world relevance of questions which employ political cartoons. "You give them a cartoon and you say, 'Interpret this cartoon,'" she said, "That's interpretation, you know? If you open a paper and you look at a picture in the newspaper and you go, 'What's that mean?' That's something you would do in real life." A middle school teacher noted she now uses DBQ kinds of questions as a regular part of her instruction:

> I was working on a social studies test today for grade seven where they have to look at a document and think about some stuff like, what was the theme about the Revolutionary war, and they've got to write notes based on the picture. And it looks—the test is a lesson. It's a lesson in analyzing documents and taking notes from the document so you're not looking to see if they're right or wrong. You're looking to see can they look and think about what's on there.

This teacher and most others praised state efforts to bring standardized assessments into closer alignment with the kind of ambitious instruction they believe is important, such as analyzing primary sources and understanding that such texts can be multiply interpreted. Social studies teachers worry about the continued strong emphasis on multiple-choice questions, but in questions like the DBQ, they see potential for pushing their students toward richer understandings.

But not all teachers held this view. Some focused on the continuing heavy presence of generally low-level multiple-choice questions, arguing that the test has changed little overall. As one middle school teacher explained:

> From my perspective, the social studies assessment doesn't seem like it's a change at all. Seems like it's kind of repackaged, kind of dressed up a little differently, but not really different and to me, there is something broken in [teachers' instruction] and we need to fix it. This new assessment to me isn't fixing it.

One might argue about whether teachers' practices are "broken," but the sentiment that some state tests, like social studies, seem less changed than others emerged throughout the focus group sessions. The English language arts and science tests, in particular, were cited as moving away from a heavy reliance on objective-style questions and toward questions with more real world and practical applications. For example, the English language arts tests asks students to write a range of pieces including technical, literary, and literary analysis essays. The science tests include performance tasks which ask students, for example, to set up a lab experiment. Teachers in these areas had questions about the nature of their respective exams, but there was a general sense that these exams push in more ambitious directions than the social studies tests do.

Social studies teachers see the prospective new state assessments as a mix of old and new. While most applaud the presence of primary sources and questions like the DBQ that ask students to analyze and synthesize information, they wonder if that emphasis won't be undercut by the continuing heavy weight of the multiple-choice section and questions which teachers generally perceive of as asking for low-level knowledge.

Opportunities to Learn About the New State Tests

New state tests, like many other educational policies, can be viewed as an occasion to learn about the craft of teaching (Cohen & Barnes, 1993; Grant, 1998b). The focus group teachers nodded in agreement when participants raised questions such as, "Do I have the skills that I need?" and made assertions such as, "We have not been taught the way we're being asked to teach....And

I think that's really difficult without a lot of staff development to get people to think differently and to teach differently."

If the need for professional development was widely expressed, the teachers' experiences suggested that they may not be getting all that they want. Studies of professional development activities suggest that what session leaders think they are "teaching" and what participating teachers think they are "learning" during professional development activities can vary dramatically (Darling-Hammond & McLaughlin, 1996; Grant, 1997a; Smylie, 1995). Consequently understanding what kinds of professional development opportunities teachers had available to them and what sense they made of those opportunities was a major element of the focus group interviews.

Three patterns emerged from analysis of the interview transcripts. One was that all teachers seemed to have had access to a wide range of professional development opportunities both around the new curriculum standards and around the new tests. A second pattern was that they found those opportunities of uncertain value. Teachers reported that the state, and occasionally district, activities often resulted in incomplete and mixed messages. The frustration many teachers expressed about the more formal professional development opportunities was mitigated, however, by their sense that working more directly with colleagues was a more profitable use of their time. The third pattern, reform by "rumor," began to emerge in the first year of interviews, but was full-blown by the second year. Despite the wide array of professional development opportunities, the teachers clearly felt that there was still much indecision about how tests would ultimately look, how they would be scored, and the like. In a context of increasing pressure to respond, but little solid information, several teachers reported the sense that rumors were driving much of their responses.

The professional development opportunities available. Asked to describe the professional development opportunities available to them, the teachers constructed a long and varied list. Some NYSED-led sessions occurred in several venues (e.g., stand-alone sessions, part of district-level inservices, sessions during professional organization conferences) and focused alternately on the new tests alone or on how the tests reflected the new state curriculum standards.

Representatives from local Board of Cooperative Extension Services (BOCES) programs also led professional development activities as stand-alone and district sessions. Some district-level sessions featured state and BOCES representatives, but others utilized the talents of district personnel, while still others brought in local and national experts. School-level professional development opportunities were also varied in that some called all teachers together, while others asked teachers to meet in grade or department-level activities. The focus group teachers also mentioned state teachers' union sessions, college and university coursework, professional literature, informal networks, and colleagues as additional sources of information on tests and testing.

The uncertain value of professional development. Of these many sources, teachers were most critical of the state-led sessions. Some felt that cuts in the NYSED have left the agency woefully understaffed. Most others, especially the high school teachers, were less generous. An English teacher said, "I'm not going to break a sweat trying to reformulate what I do when their people (NYSED) don't know what they're doing." A social studies teacher was more blunt: "Do they have a clue as to what's going on?"

District-level sessions received more mixed reviews. A high school mathematics teacher praised her district's efforts to develop professional development activities that would meet teachers' perceived needs:

> My district is real supportive. If I say to them we need an inservice on blah, they will say we'll do it. They're wonderful that way. It's very teacher driven. Our school district is wonderful as far as them involving teachers and listening to the teachers and valuing what the teachers say.

This comment stood largely alone, however, as most other teachers suggested that district-led professional development was lacking in usefulness. A high school social studies teacher noted:

> We've had two district wide superintendents conference days and we've talked about [the tests] and gone over some things, but not into the detail that needs to be done to get a good

feel for the types of questions and changes. I think in our building many people would still be hard pressed to give an accurate reflection of what the assessment is all about.

A middle school science teacher attended a district-sponsored inservice led by a district teacher. She reported that while session could have been valuable, she left frustrated because the teacher who led the session came from a magnet school where resources are plentiful whereas she teaches in a resource-starved neighborhood school. Not all the blame for weak district-sponsored professional development was laid at the feet of the leaders, however. A secondary social studies teacher panned the district-level sessions she attended, but she assigned much of that responsibility to her colleagues:

> We went to the district-wide [inservices]. They (the inservice leaders) always tried to be very positive, but the overwhelming number of teachers who are so negative about this assessment always wins out. It basically becomes a complaining session and you really aren't focusing on what the whole meeting was about anyways.

The focus group teachers reported that school-, grade-, and/or department-level professional development activities were generally more useful than state or district efforts. An elementary school teacher, for example, praised the work her grade-level colleagues were doing:

> We have grade-level meetings. They're very positive, you know, even though we all don't want to test, we all feel like we shouldn't have to do it. They're (her colleagues) always very positive, always very friendly approaching it. Every time we go to a grade level meeting, [the team leader] always is handing us stacks and stacks of information materials. Things that we might need or might be able to use to help the kids get ready, whether it's for the science or the math or the English [tests]. There's always something positive going on.

A high school mathematics teacher explained that not only has the amount of conversation increased in her department, but that it is becoming increasingly acceptable to say, "I don't know how to do this." She went on to describe how her colleagues, both veteran and novice, were creating a new ethic whereby the traditional norms of isolation and "doing your own thing" were fading.

Not all teachers are similarly situated, however, and more than any other group, the high school social studies teachers present described their departmental interactions as less than optimal. Several nodded in agreement when an untenured teacher portrayed her colleagues as being obsessed with talk about "how to beat the test, or change the test, or fight the state, or fix the state or...how is the administration wrong, how are we right." Potentially useful discussions of teaching, learning, and assessment, she explained, get lost in the mix.

If teachers found formal state, district, and school-level professional development of uncertain value, all reported instances where informal networks and relationships had proven valuable. A high school social studies teacher said that, while she appreciated some elements of her district staff development days, "it is a lot easier to bounce off the ideas with somebody. And I just wrote [a DBQ] a few weeks ago with a colleague. We have now the same planning period so that worked out." A high school teacher reported that she and her colleagues have met informally after school to consider assessment issues. "There were a handful of us that got together after school on a voluntary basis," she said, "....It makes my life a lot easier when I talk to other English teachers." In addition to these unstructured activities, several elementary school and high school mathematics teachers described informal networks of educators who meet regularly to discuss a range of issues, including those related to testing. A mathematics teacher described the benefits she has appreciated from her involvement:

> We have each other (she laughs). We have a network through (a local state university)...where there have to be what—about 70 teachers, maybe 100 maybe that—we have meetings four times a year, and so now I don't feel isolated anymore. I mean I can always call [a colleague in a neighboring district]. I have friends [in another district]. Friends

just about anywhere. I know what's going on at what school and I can pool resources, and so that helps a lot.

The power of such informal relationships is apparent: These teachers sense that they are working with peers who hold similar goals and concerns, who are willing to share ideas and practices, and who offer a sense of belonging. Such relationships, then, have an immediacy and a specificity that seems missing from the more formal professional development opportunities teachers typically experience. That these teachers have sought out and participated in these relationships is admirable; that they have felt compelled to do so in order to meet their needs is ironic, however, given the seeming wealth of structured opportunities.

Reform by rumor. Having informal sources of information and support may help teachers navigate some of the challenges the new state tests posed, but they do little to help teachers with the problems of mixed messages and unanswered questions. In fact, the more sources of information teachers encounter, the greater the incidence of reform by rumor.

Common across teachers of all grade levels and subject matters was a frustration with incomplete and conflicting information about the new tests. An elementary school teacher noted, "If we just had more information and if we knew what was expected of us and how to do it, possibly…we could do what was expected of us." A high school mathematics teacher added:

> If they're (NYSED) going to give us information, they have to give it more structured backing. Not this haphazard changing the rules daily….Our math department head has said [at an inservice led by an NYSED representative], "Tell us what you want. We will do it. We will change the way we teach….But you can't keep changing the messages you're giving us."

To be sure, state leaders seem to recognize that they are sending multiple and, at times, confusing messages. A high school mathematics teacher reported the following experience during a state-sponsored inservice:

> When we go to state meetings, (the NYSED representative) who's in the math ed department always prefaces his remarks with, "What I'm going to tell you is true at May 13th at 4 whatever. It's true right now. When I go back to my office, it might not be true." And we

get to go to a lot of state meetings and everything and find out what's going on. And we always find out the latest stuff, but then it changes.

As this quote suggests, teachers do not necessary blame the state education representatives, but they are frustrated with the uncertainty of the situation. A high school social studies teacher's experience summed up some of the anxiety mixed and multiple messages can induce:

> I don't know if this geography thing (i.e., that the state curriculum and test for tenth grade were changed from Global Studies to Global History and Geography) is true or not. But somebody in my department had been in the state conference the week before and said, "I didn't hear any of this." And then we started frantically calling—I think we called the (local state university) Social Studies department, and they were calling all over to find if this was true. And I think the final verdict was that, "yes (geography has been added), but geography the way we've always taught it, so don't be nervous. They (NYSED) are not asking to name which direction the Danube River flows or anything like that." But, I don't know. It's crazy.

This teacher went on to remark, "I see it as just lots of rumors. It's like every other day we're coming in, 'Did you hear they're cutting out the constructed response? Oh, now the new course is Global History and Geography?'"

A cynical interpretation of the above is that teachers are merely pawns in a game that is being transacted all around them. This view asserts that while changing teachers' practices is the target, teachers' ideas and voices are largely ignored as those above them—state and district-level actors—do the real work of policy change. Teachers, through their professional development opportunities, may listen in. But as listeners rather than as full participants, they hear only bits and pieces, and rumors rule the day.

A more generous interpretation has two elements. One is that reforming education is simply hard work, especially when done in midstream, or what a policymaker in another state termed, "rebuilding the airplane while you're flying it" (Lusi, 1997, p. 91). The second element is that, given the sheer number of teachers and the wide range of circumstances in which they work, policymakers face a daunting task in attempting to change pedagogical practices. Whether they

should try to or not, the parameters of the NYSED operation are intimidating: thousands of teachers, in thousands of schools, in close to 700 districts, and an agency with little more than a handful of employees. Clearly, then, NYSED must rely on the efforts of proxies—BOCES educators, professional organizations, district and school-level leaders, college and university academics—who may or may not understand and/or support the state agenda. In such a situation, the potential grows for mixed and confusing messages, and for reform by rumor.

The Rationales for and the Consequences of the New NYS Tests

The notion of "reform by rumor" functioned as a proxy for a number of comments where focus group teachers talked about feeling left out of the conversation about changing state assessments. Teachers across grade levels and school subjects expressed frustration that, while they are the professionals on whom the tests will have the most impact, their voices are not well reflected in important discussions about the nature, import, and design of new state tests. As one teacher said, "I really fear that unless there's open communication...this whole thing would be just kind of a charade." Another added, "I just feel that I've been talked at."

These teachers remain uncertain about the rationales for and the consequences of the state assessments, but seek to question rather than condemn. Most said they have attended meetings designed to inform them about the tests, but none said they were satisfied: Their questions either went unaddressed or, if they were addressed, the information they received did not always jive with information circulated previously. While numerous questions arose during the focus group interviews, two dominated: questions about the rationales for changing the assessments and questions about the intended and unintended consequences of the tests.

Questioning the rationales for the tests. Whether the NYSED hopes to induce changes in teachers' curriculum decisions, their instructional practices, or both has been unclear for some time (Grant, 1997a). The focus group teachers echoed this confusion. They also discussed their uncertainty about whether the state's intention was change their behavior or the students'. As a middle school social studies teacher said, "Are they (NYSED) doing this to better students' education, or are they doing it so they can say, 'Look, we changed something.'"

On the question of whose behavior NYSED is targeting, teachers expressed considerable frustration. For instance, an elementary teacher asked, "Who is it assessing? Is it really assessing the students? Or is it assessing the teachers?" Another elementary teacher echoed this point: "What is the purpose of the state exams? Is it actually to assess the students or to push the teachers in a direction?" A secondary social studies teacher spoke directly to the issue of whose life is changing the most as a result of the new state tests:

> I think it's ironic that the state came out with all of these decisions in order to improve student learning and to make students better students and...I feel like I am doing so much work this year. When I do essays, I try to fix things and give them lots of responses and they just—I feel like I'm doing more work than the kids sometimes....The last couple weeks it's like "I'm not taking this test! I took this test!" This is you. Not me. But it seems like the teachers are on the chopping block. And it's just ironic that it's no longer the student anymore. And it's the kids who are taking the test. And it seems like the kids are almost less and less responsible....

The last part of the quote above suggests that the issue of whether teachers or students are targeted is important, in part, because teachers are unsure where the blame is going to come down should test scores not rise. Many suspect, however, that teachers will take the brunt of the criticism. A high school mathematics teacher said, "They're (local administrators) are going to be pointing their finger if your kids don't do well. They're going to be pointing their finger at those teachers and that's unfortunate because they're (the teachers) going to be a scapegoat because of it." A secondary English teacher talked about the unfairness of holding the teachers whose students are taking the tests entirely responsible for the outcomes:

> I think that whole culture needs to change because you are not the sole responsible party for that student's abilities....If someone did a lousy job last year, then you're getting a group of students without the proper foundation. And is there going to be some kind of mechanism that will address that if you realize that the child did not get proper foundation?...There's no way I solely am responsible for that child's [test scores]. I've had students who are

functioning very very low and you're asking me to...bring that child further along. Is that child going to pass that test? No. So you're going to come to me and say, "Well, only 55% of your students passed this test. You're lousy!" I'm going to say, "Well, what did you give me?"

This quote raises a number of thorny issues, not the least of which is a seeming deficit view of children. This view implies that students come to a teacher with a set of deficiencies, resulting from poor parenting, poor schooling, and the like, which the teacher must then "correct." The problems with this view are several, but in this case, they serve to amplify the dilemma this teacher faces: She feels the twin burdens of preparing students to take the exam and of being held accountable for their performance. Although it seems unfair to make the child the pawn, this teacher rightly points out that she alone can not be responsible for test scores.

Teacher frustration was also apparent around the question of whether NYSED's intent was to change curriculum, instruction, or both. The focus teachers assumed the tests were meant to induce changes, but they were unsure what sort of change was expected.

A secondary social studies teacher saw the state's aim as primarily directed toward curriculum:

But it looks like -- the more I hear about it it's as if the state through its tests is controlling what gets taught in the classroom. By saying that the test is going to be done this way, all of a sudden it's going in and saying well you can't teach this, this, and this when you want to. You have to teach this. You have to teach this.

An elementary teacher, by contrast, suspected that the state's intention is to influence teachers' instructional practices:

Is this a way of making teachers look at their practice and alter their teaching techniques because they see a certain topic being covered on an exam and so they'll say, "Oh, I didn't do that so well that time. I guess I have to spend more time on that next year." So if you see the focus on the exams, then you've got to go back and make sure that you include that type of

instruction the next year. And so I think—are the tests pushing—is the state using the test to push teachers in a certain direction with their instruction?

While most of the focus groups sensed that the state tests were being used to leverage change of one sort or another, not all did. A high school English teacher reported that she had been told, "We've been doing this all along. That this is no big deal...all we have to do is get kids accustomed to the format [of the test]." A secondary science teacher added to this notion, by reciting a familiar teacher expression, that is, "this too shall pass." "In our science department," he said, "they feel because science is the last assessment [to be introduced] that this is all going to blow over." The notion that whatever NYSED introduces is likely to fade in importance over time was not the dominant view among the focus group teachers. But it's expression should warn state-level reformers that whatever leverage they believe tests hold for changing instruction and/or curriculum may be illusory. This is not because teachers do not sense that problems exist: None of the focus group teachers was willing to suggest that all is right with public education. But several supported the following sentiments of an elementary school teacher who questioned the reliance on tests as lever of real instructional change:

> I understand that certainly there are places in American education that are in dire need of shaping up somehow....It (the test) just seems to me a misdirection of resources. We're spending how much--thousands of dollars on training, on writing these tests or whatever they're doing to when the real issue is what's happening in the classroom. What kind of preparation are teachers getting? What kind of preparation are they getting before they even get a classroom? What kind of thinking is going on here? And are those questions even being asked? Or were they ever asked before this happened? It was just suddenly that we had this massive assessment. And I don't remember any sort of input from teachers. I don't remember any state education people coming to us and saying, "What do you think?" Or, "What's going on in your classroom?" It was just this kind of mandated attempt to reform. And maybe it will work. I mean, I don't know whether it will work or not. But it seems to

me there's so much more that could be done that hasn't been attempted in terms of helping teachers.

To be fair, NYSED officials and the state Board of Regents have proposed a range of reforms that push changes in curriculum and in teacher education. The primacy of the state testing program, however, weighs heavily. The focus group teachers were not opposed to improving teaching and learning, but they are uncertain about the rationale for standardized tests as a vehicle.

Predicting the consequences of the new tests. The idea that the new tests may yield no real consequences for teachers' practices was one of several predictions the focus group teachers made. Most of those predicted consequences were negative, but not all. For example, several teachers in the first year focus groups expressed the hope that the tests would mean greater collaboration with their colleagues. A high school English teacher summed up the feeling: "If there were more opportunities to get more people together, that would help." While it was far from unanimous, a number of the year two teachers reported that, in fact, they had found their peers receptive to and interested in working together.

The overwhelming sentiment, however, was that the new tests could produce undesirable effects. Those effects grouped loosely around issues of pedagogy, students, and teachers.

Two related consequences of tests for pedagogy arose. One is that, rather than promote more ambitious teaching and learning, the state tests may actually push more reductive forms of teaching and learning. The most common expression was that teachers felt increased pressure to tailor one's teaching to the test parameters. As a secondary social studies teacher noted, "You've got people in high places just saying 'teach to the test.'" A middle school English teacher complained that he felt pressure to "teach them (students) test terminology when I could be teaching them other things." This teacher went on to describe the kind of support his district provides as little more than practice exercises. "The only thing I've gotten from my district," he said, "is lots of practices. Every week there's, 'Thank so and so for giving this practice material. Here's another listening practice that you may want to use.' I could have spent my whole year doing practices."

The sense that teachers feel pressed to adopt direct teaching approaches as a means of bolstering short-term test performance was in direct competition with the sentiments expressed earlier that the new state tests could be viewed as supportive of more ambitious instruction. During the interviews, however, no teacher commented on this seeming contradiction. One explanation is that they were simply unaware of its emergence. A more interesting possibility is that these teachers can read multiple messages in the tests. Take social studies as an example. Teachers thinking about the multiple-choice questions could reasonably assume that a more traditional, direct instruction approach was being encouraged. If those same teachers were thinking instead about the DBQ questions, it seems equally reasonable to assume that richer forms of pedagogy were intended. This ambivalence, which has surfaced in a number of places already, underscores the difficulty in understanding teachers' perceptions of state tests and it suggests that their classroom responses may be more complex and textured than reformers may want or expect.

A second potentially negative consequence of the new tests was an increased emphasis on remediation as a way to deal with low test scores. The teachers, especially those in the second year interviews, described a wide array of remedial approaches taken in their schools. Those approaches included additional classes designed for students presumably at risk of failing, summer and Saturday test review courses, hiring additional teachers and aides to staff learning labs where students could either come voluntarily or by teacher assignment, and reassigning teachers to classes of students based on their perceived ability to help those students pass the exam.

The teachers offering these examples generally seemed supportive of them. The seeming contradiction that ratcheting up remedial efforts would occur at the same time teachers were being pushed to change their pedagogy went unremarked upon. Again, however, this contradiction may be less apparent than one might suspect. Empirical evidence is surprisingly thin on the question of which instructional approaches lead directly to high test scores (Cohen & Barnes, 1993; Grant, 1998a). Consequently, a reasonable response to a new testing situation might be both to make changes in "regular" classes and to begin planning for remedial instruction at the same time.

The real danger, however, is that these remedial opportunities will become little more than drill sessions, a point that was recognized by several teachers. For example, a high school mathematics teacher observed:

> If the students do not pass, they're going to be remedied with questions that will make them pass. So eventually every student will pass. Doesn't matter the categories, they're going to do component retesting, so if the student doesn't do well in these three areas, they'll be grilled in those three areas with a bank of questions, and then the student will have another test from the bank that he was drilled in. So eventually they'll get it.

Such an approach may work for low-level skills, but is of dubious use in areas like social studies where conceptual knowledge is central. As VanSledright & Brophy (1992) observe, "naïve but imaginative accounts persisted in some children even after direct instruction designed to change them" (p. 854). Without any definitive research supporting one means of improving test performance over another, drill and practice remediation is as likely to flourish as any other approach.

A second area of negative consequences anticipated by the focus group teachers concerned students. An elementary teacher worried generally about that the net effect of a high profile, high-stakes testing program would be a "nation of test-takers":

> Something that I've been thinking about more is the effect this has on the children, on the student. What kind of learners is this going to shape? Are we producing a nation of test-takers, and if so, are those test-taking techniques or skills what we need to produce life long learners that we talked about before?

Other teachers expressed more focused concern about the anticipated consequences for urban students. Wiles (1996) argues that test performance is clearly distributed along socio-economic lines with upscale, white suburban children consistently outscoring their urban and minority peers. The focus group teachers, both urban- and suburban-based, recognized the inherent threat that high-stakes testing poses for some children. An elementary school teacher said, "I'm very concerned about some of the larger populations in the bigger urban areas. I don't understand how

this is going to positively affect these kids." A high school teacher, commenting on the anticipated testing of special education students, asked, "How do we accommodate the non-standard kids on a standardized test?"

No teachers thought their students' scores on the new tests would improve over past test scores. A couple of teachers did express, however, the hope that their students' scores would increase over time. A middle school English teacher said, "I think, naive though it may be, that our kids are going to do better ultimately on these exams. Maybe not this year, but ultimately."

This hopefulness stood in stark contrast with the prevailing view that teachers anticipated problems for their students. Underlying both these sentiments is a harsh truth: These teachers simply do not know how their students will perform on the new tests. Given the general tendency for a correlation between test scores and students' social capital, it is difficult to understand why suburban teachers would be worried. And yet, analysis of the relative concern expressed by suburban vs. urban teachers suggested that suburban teachers and administrators may be even more concerned about potentially low scores than their urban peers. One proxy for this finding is the observation that the overwhelming number of remedial efforts planned are being developed in suburban schools.

As noted above, no teacher feels s/he has an inside track on what approaches will insure high scores. Left to follow one's hunches, it is no particular surprise to find concern among all teachers, both suburban and urban. But what explains the fact that suburban teachers seem to be *more* concerned about their students' performance than their urban peers? Part of an explanation must consider the notion that not all suburban districts are created equal. The suburban teachers in focus group teachers represented first-, second-, and third-ring suburbs. First-ring suburbs tend to include a range of working to middle class students. Second-ring suburbs are more upscale; most students come from middle to upper-middle class homes. Finally, the third-ring suburbs are rural areas that recently have attracted a large number of middle and high SES families. With the exception of one or two urban magnet schools, it is the schools in the second- and third-ring suburbs that consistently rank in the top quartile according to a highly publicized local business

magazine. Top quartile spots on this list have real consequences for real estate values, bragging rights, and the like, and so the scramble to move up can be intense. New tests, then, represent a potential threat to schools' past standings. School people in high performing schools want to maintain their positions; educators in middle and low performing schools hope to at least avoid dropping further.

The competition for high test scores plays out as a third set of consequences. Here, the focus is on the pressure and uncertainty teachers feel as they decide if and how to modify their teaching based on their perceptions of the state test. A couple of these pressures have already been described. One is the feeling of uncertainty teachers have about which approaches will ensure higher scores. A second pressure surfaces as teachers report being made to feel entirely responsible for their students' results. Putting the point on this feeling is a secondary social studies teacher:

> Just this week I was called down to the office and we were comparing some of the *Business First* statistics that were out just recently....So according to our administration [if we get low test scores]...people come out to vote and decide they don't want to vote on the budget, therefore the whole community goes down. So, I left the office thinking the weight of this town...is on my shoulders. Whether or not, you know, my kids pass. And we had like a 70% last year and we're expected to have at least a 90 if not higher. So, in terms of administration, testing is a pretty big deal.

Not all principals apply pressure so directly, but many apparently do. This is more likely to happen in high schools than elementary schools, however. According to several of the focus group elementary school teachers, their principals are more likely to talk about test scores as part of bigger picture of how students are progressing. These teachers do not necessarily feel any less pressure than their high school peers, but one source of pressure, the school administrator, seems to be less of a factor.

The new elementary school exams are more high-stakes than they used to be; recall that now individual student scores will be reported rather than group scores. The stakes are even higher in

the high schools, however, as passing the Regents exams will be necessary in order to graduate. Consequently, it is not hard to understand why high school administrators might be more likely than their elementary peers to put pressure on their teachers. Whether that tactic will pay off ultimately or not is hard to predict. But one manifestation of that pressure is to cause teachers to consider issues that they probably have not had to think about in the past. One particularly compelling story came from a high school social studies teacher who said she now wonders about each new student who comes into her classes:

> I never--it never crossed my mind before that a certain kid was going to lower my passing rate or not, and I actually started thinking about that this year. And I was so ashamed of myself about that. And one of the girls I had transferred from a general track. She stayed in my class. I didn't want to just dump her. But she can now take the RCT at the end of the year. But I had a girl a couple years ago who transferred from another state. She never had Global 9. And I was just happy to work with her and she was going to try it. And if you go to look at an individual kid and say they're not going to do it, it's horrible to think that--to individualize it like that. Because I guess every couple kids knocks you down a little bit. And our--I know that our department chairs had our results individualized and our principal keeps coming into meetings saying, "How can we raise this up? How can we do this better?"

This teacher concluded her story with a nervous laugh, saying, "But I'm glad I have tenure, right?" Yet, having tenure seems little consolation for this thoughtful and dedicated teacher now confronted with the dilemma of wanting to work with all students, but recognizing that doing so may cause her teaching to be called into question should her students' scores not measure up.

Not all the consequences described were negative, however. Several teachers cited greater collaboration with their peers as a key benefit of the new tests. Elementary teachers and high school mathematics and English teachers were most vocal on this point. "I think we have so much to learn from each other," one elementary teacher said. Another echoed this point, commenting, "We're really trying to deal with this [new tests] and trying to work as a faculty to help each other." A high school English teacher noted that information is vital and that colleagues are an

important source, "What's most important to me is being able to communicate with other people so I can get some information." A high school mathematics teacher concurred, but pointed out that that the new exams were forcing teachers to rely on each other:

> I think the nature of the testing--it certainly sets the situation up for teachers to talk. Because the types of questions that happen to be asked. They don't have the stockpile of old Regents questions. So [teachers say] "I came up with this. You know, I'm going to use this." We can share, and the nature of the beast is forcing the issue.

Social studies teachers reported some positive collaborations with peers, but they also cited more instances than the other teachers of situations where friction had developed. A high school teacher described the tension that arose over course assignments:

> We have attempted to get together and work, but what we have found out has been happening is just been a lot of back-stabbing and a lot of animosity because there are a couple of teachers who just adamantly refuse to teach 10th grade (when the Global exam is administered). So the feeling is, well, they can do the ninth grade program. But where is their accountability? Because they just will not do that 10th grade when their kids take the Regents at the end of the year.

This teacher's experience points, again, to the variability in the way consequences of the test are playing out. This variation is explained, in part, by the development of as many unintended as intended consequences. State-level reformers may have hoped, for example, that teachers would see the test as an impetus for more ambitious instruction, closer collaboration, and the like. And this seems to be occurring. But reformers probably did not predict the more negative consequences these teachers are seeing. That these outcomes are unintended is little solace, for they may be just as real to the teachers as the intended outcomes. Actually, these unintended consequences may ultimately be more important because they seem to receive scant attention from state and district -level actors. State and district leaders may be unaware of these issues, they may be ignoring them, or they may not see them as problems. In any event, it seems interesting that no teacher mentioned that s/he had participated in any explicit conversations about the problems they

anticipated. As noted above, teachers did see positive possibilities arising from the new state tests and there was no particular sense of gloom during the interviews. How teachers will manage the more negative consequences is unclear, but the supposition that they will have no effect seems naïve.

Implications

Substantive change is always unsettling. So reform on the scale that New York state is attempting, in all grades and in all school subjects, is bound to generate some frustration, anxiety, and uncertainty. The findings above tell us that while teachers are not adverse to change, they have real concerns about the nature of the changes proposed, the professional development opportunities available to learn about these changes, and the rationales for and consequences of the new state tests.

Given the complexities of teaching and policy (Grant, 1998a), it is not surprising to learn that teachers see both prospects and problems in the new NYS tests. What this study suggests, however, is that teachers are not passive participants and must not be designed around. The dream of teacher-proof curriculum as a means of changing teachers' practices has proven to be a myth (See, for example, Dow, 1991; Schwille et al., 1983). Faith in tests as a means of corralling teachers' practices may ultimately prove just as chimerical as long as teachers are left out of the loop. If any of the changes state reformers propose are to stick, then these teachers are saying they need to be more actively involved in the formulation of those changes. But there is something else. These findings also suggest that there are real and important differences in the ways teachers perceive reforms across grade levels. Among other things, this means that reformers can not take a one-size-fits-all stance and that professional development needs to be sensitive to the differences in the perceived needs of teachers.

References

Bogdan, R., & Biklen, S. (1982). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn and Bacon.

Cohen, D., & Barnes, C. (1993). Pedagogy and policy. In D. Cohen, M. McLaughlin, & J. Talbert (Eds.), *Teaching for understanding: Challenges for policy and practice* (pp. 207-239). San Francisco: Jossey-Bass.

Comfort, K. (1991). A national standing ovation for the new performance testing. In G. Kulm & S. Malcolm (Eds.), *Science assessment in the service of reform* (pp. 149-161). Washington, DC: American Association for the Advancement of Science.

Corbett, H. D., & Wilson, B. (1991). *Testing, reform, and rebellion.* Norwood, NJ: Ablex.

Darling-Hammond, L., & McLaughlin, M. (1996). Policies and support professional development in an era of reform. In M. McLaughlin & I. Oberman (Eds.), *Teacher learning: New policies new practices* (pp. 202-219). New York: Teachers College Press.

Dow, P. (1991). *Schoolhouse politics: Lessons from the Sputnik era*. Cambridge, MA: Harvard University Press.

Editors. (1994). Introduction. Symposium: Equity and educational assessment. *Harvard Educational Review*, *64*, 1-3.

English, F. (1980). *Improving curriculum management in the schools*. Washington, DC: Council for Basic Education.

Feltovich, P., Spiro, R., & Coulson, R. (1993). Learning, teaching, and testing for complex conceptual understanding. In N. Frederikson, R. Mislevy, & I. Bejarq (Eds.), *Test theory for a new generation of tests* (pp. 181-217). Hillsdale, NJ: Lawrence Erlbaum.

Finn, C. (1995). Who's afraid of the big, bad test? In D. Ravitch (Ed.), *Debating the future of American education: Do we need national standards and assessments?* (pp. 120-144). Washington, DC: Brookings Institution.

Firestone, W., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, *20*(2), 95-113.

Freeman, D., Kuhs, T., Porter, A., Knappen, L., Floden, R., Schmidt, W., & Schwille, J. (1980). *The fourth grade mathematics curriculum as inferred from textbooks and tests* . East Lansing, MI: Institute for Research on Teaching, Michigan State University.

Fuhrman, S. (1993). The politics of coherence. In S. Fuhrman (Ed.), *Designing coherent education policy: Improving the system* (pp. 1-34). San Francisco: Jossey Bass.

Fuhrman, S., Clune, W., & Elmore, R. (1988). Research on educational reforms: Lessons on the implementation of policy. *Teachers College Record*, *90*, 237-258.

Glaser, B. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.

Glatthorn, A. (1987). *Curriculum leadership*. Glenview, IL: Scott, Foresman.

Grant, S. G. (1996). Locating authority over content and pedagogy: Cross-current influences on teachers' thinking and practice. *Theory and Research in Social Education*, *24*(3), 237-272.

Grant, S. G. (1997a). Opportunities lost: Teachers learning about the New York state social studies framework. *Theory and Research in Social Education*, *25*(3), 259-287.

Grant, S. G. (1997b). A policy at odds with itself: The tension between constructivist and traditional views in the New York state social studies framework. *Journal of Curriculum and Supervision*, *13*(1), 92-113.

Grant, S. G. (1998a). *Reforming reading, writing, and mathematics: Teachers' responses and the prospects for systemic reform*. Mahwah, NJ: Lawrence Erlbaum.

Grant, S. G. (1998b). *An uncertain lever: The influence of state-level testing on teaching social studies*. Paper presented at the American Educational Research Association, San Diego, CA.

Heubert, J., & Hauser, R. (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.

Kellaghan, T., Madaus, G., & Airasian, P. (1982). *The effects of standardized testing*. Boston: Kluwer-Nijhoff.

Koretz, D. (1988). Arriving in Lake Woebegon: Are standardized tests exaggerating achievement and distorting instruction? *American Educator, 12*(2), 8-15; 46-52.

Koretz, D. (1995). Sometimes a cigar is only a cigar, and often a test is only a test. In D. Ravitch (Ed.), *Debating the future of American education: Do we need national standards and assessments?* (pp. 154-166). Washington, DC: Brookings Institution.

LeCompte, M., Preissle, J., & Tesch, R. (1993). *Ethnography and qualitative design in educational research*. (2nd ed.). New York: Academic Press.

LeMahieu, P. (1984). The effects on achievement and instructional content of a program of student monitoring through frequent testing. *Educational Evaluation and Policy Analysis, 6*(2), 175-187.

Lusi, S. F. (1997). *The role of state departments of education in complex school reform*. New York: Teachers College Press.

Madaus, G. (1988). The influence of testing on the curriculum. In L. Tanner (Ed.), *Critical issues in curriculum: 87th yearbook of the NSSE, Part 1* . Chicago: University of Chicago Press.

National Center for Education Statistics. (1999). *Teacher quality: A report on teacher preparation and qualifications of public school teachers*. Washington, DC: Author.

National Center for History in the Schools. (1994). *National standards for United States history*. Los Angeles: Author.

National Council for Social Studies. (1994). *Expectations of excellence*. Washington, DC: Author.

Natriello, G., & Pallas, A. (1998). The development and impact of high stakes testing. Available on-line: http://www.columbia.edu/~gjn6/histake.html

New York State Education Department. (1996). *Learning standards for social studies*. Albany, NY: Author.

New York State Education Department. (1998). *Social studies resource guide*. Albany, NY: Author.

New York State Education Department. (1999). *Global history and geography Regents examination: Test sampler draft*. Albany, NY: Author.

Popham, J., Cruse, K., Rankin, S., Sandifer, P., & Williams, P. (1985). Measurement-driven instruction: It's on the road. *Phi Delta Kappan, 66*(9), 628-634.

Popham, W. J. (1998). Farewell, curriculum: Confessions of an assessment convert. *Phi Delta Kappan*, , 380-384.

Ravitch, D. (Ed.). (1995). *Debating the future of American education: Do we need national standards and assessments?* Washington, DC: Brookings Institution.

Resnick, D., & Resnick, L. (1985). Standards, curriculum, and performance: A historical and comparative perspective. *Educational Researcher, 14*(4), 5-20.

Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? *Phi Delta Kappan, 62*, 631-634.

Schwille, J., Porter, A., Belli, G., Floden, R., Freeman, D., & Knappen, L. (1983). Teachers as policy brokers in the content of elementary school mathematics. In L. Shulman & G. Sykes (Eds.), *Handbook of teaching and policy* (pp. 370-391). New York: Longman.

Shanker, A. (1995). The case for high stakes and real consequences. In D. Ravitch (Ed.), *Debating the future of American education: Do we need national standards and assessments?* (pp. 145-153). Washington, DC: Brookings Institution.

Smith, M., & O'Day, J. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 233-267). New York: Falmer.

Smylie, M. (1995). Teacher learning in the workplace; Implications for school reform. In T. Guskey & M. Huberman (Eds.), *Professional development in education: New paradigms and practices* (pp. 92-113). New York: Teachers College Press.

Stiggins, R., & Conklin, N. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. Albany, NY: SUNY Press.

Wiles, D. (1996). *Networking high performance in New York's secondary education: The Regents curriculum story*. Lanham, MD: University Press of America.

Wolf, R. (1998). National standards: Do we need them? *Educational Researcher, 27*, 22-5.

Appendix

FOCUS GROUP PROTOCOL
Spring, 1998

Introduction: Why we are here. Guidelines and ground rules.

METAPHORS
Moderators and participants introduce selves to group.

To get started introduce yourself to someone next to you and describe an image or metaphor that characterizes your thinking and/or feelings about the new state assessments.

After they have shared in pairs, have them share their metaphors with the group.

Have participants discuss and elaborate on the metaphors. Lead a discussion of the metaphors. What do they say about our thinking? Common features? Significant differences.

Direct the discussion toward the next question—what do these assessments mean to you?.

MEANING OF ASSESSMENTS
What do/will these assessments mean to you? Your school? Your students?

Transition to next question—are you prepared to deal with these implications?

BEING PREPARED
How prepared to deal with these assessments do you feel? How are you being prepared? What are you being prepared for? What opportunities do you have to talk about the assessments and related issues?

Out of the discussion to this point we should have expressions of frustration, dissatisfaction, confidence, disdain. Build on these expressions to move toward a discussion of needs.

What help do you need?

This discussion should lead naturally to talk of challenges.

CHALLENGES
What challenges/concerns do you anticipate? How will you deal with these challenges/concerns? Who do you expect will help you?

CLOSURE
What has this conversation made you think about concerning teaching and testing (e.g., issue, question, new image) ?

Next step—feedback form. Thanks for coming!

---

[1] Corbett and Wilson (1991) point out, however, that Madaus' claims are based on limited data: "anecdotes, testimony from public hearings, historical accounts, and an occasional international study" (p. 26).
[2] Revisions of state tests is still in progress so some of what follows is based on SED reports of changes they expect will occur.
[3] The first administrations of new social studies tests will begin in the fall of 2000.

[4] For example, in the test sampler for the Global History and Geography exam (New York State Education Department, 1999), students would be given documents that range from a poem by Lao Tzu; portions from Pericles' "Funeral Oration," the English Bill of Rights, the Japanese Constitution, a speech by Benito Mussolini; and a political cartoon about the monarchy in France during the 1600-1700s. They are then directed to write an essay in which they "compare and contrast the different viewpoints societies have held about the process of governmental decision making and about the role of citizens in the political decision-making process" and to "discuss the advantages and disadvantages of a political system that is under the absolute control of a single individual or a few individuals, or a political system that is a democracy" (p. 25).

A test sampler in NYS consists of a description the types of test items, sample questions, a breakdown of the number of questions by curriculum standard and topic, rubrics for essay questions, and sample student responses.

At present, the only test sampler available is that for tenth grade Global History and Geography. The first administration of that test is scheduled for June 2000. Test samplers for the grades 5 and 8 tests are to be available this fall with administration of the grade 5 test scheduled in November 200 and the grade 8 test in June 2001. The test sampler for the grade 11 test is due out in spring 2000 and the new test is scheduled for June 2001.

[5] From the Global History test sampler (New York State Education Department, 1999), students are given this theme on belief systems: "At various times in global history, members of different religions have acted to bring people together. Members of these same religions have also acted to divide people and have caused conflict." Students are then directed to this task: "Choose *two* religions from your study of global history and geography. For *each* religion: Describe *two* basic beliefs of the religion; Explain how members of the religion, at a specific time and place, acted either to unify society or to cause conflict in society" (p. 29).

[6] The PET tests were given at grades 6 and 8. The new tests will be administered at grades 5 and 8.